HARVARD | BUSINESS | SCHOOL

Discrete Within the Constraints of Opportunity: Gender Homophily and Structure in a Formal Organization

Adam M. Kleinbaum
Toby E. Stuart
Michael L. Tushman

Working Paper

12-050

December 21, 2011

Discretion Within the Constraints of Opportunity: Gender Homophily and Structure in a Formal Organization[•]

Adam M. Kleinbaum
Dartmouth College
Tuck School of Business
adam.m.kleinbaum@tuck.dartmouth.edu
603.646.6447

Toby E. Stuart
Harvard Business School
tstuart@hbs.edu
617.496.4626

Michael L. Tushman
Harvard Business School
mtushman@hbs.edu
617.495.5442

Working Draft

August 23, 2011

[•] The authors are grateful to Pino Audia, Pierre Azoulay, Matt Bothner, Connie Helfat, Andy King, Gail McGuire, Ray Reagans, Marc-David Seidel, Judith White, and Valery Yakubovich; seminar participants at Harvard Business School; and conference attendees at the Academy of Management for helpful comments and suggestions and especially to Paul Wolfson for helpful suggestions and valuable statistical and programming support. The first author gratefully acknowledges financial support from the Ewing Marion Kauffman Foundation. Please direct correspondence to Adam M. Kleinbaum.

Discretion Within the Constraints of Opportunity:
Gender Homophily and Structure in a Formal Organization

Abstract

Homophily in social relations is widely documented. We know that homophily results from both individual preferences and uneven opportunities for interaction, but how these two mechanisms interact in formal organizations is not well understood. We argue that organizational structures and geography delimit opportunities for interaction, but that within the opportunity sets created by business units, job functions and offices, actors have a greater level of discretion to choose their interaction partners. Therefore, we expect to observe more homophilous interactions within these structures than across their boundaries. We test this argument using a dataset consisting of millions of e-mails exchanged among thousands of employees in a large information technology firm. We find significant interaction effects between being of the same sex and being in the same business unit or same office on dyadic communication rates, though not with same job function. In an extension, we find that men's communication patterns are consistent with this theory, but that women communicate differently: relative to male-male and male-female pairings, female-female interactions are much more likely to occur across organizational boundaries. These findings have implications for research on homophily, gender, and formal and informal structure in organizations.

Social interactions are notoriously homophilous (McPherson, Smith-Lovin and Cook 2001). Across a range of relationship types (e.g., Lazarsfeld and Merton 1954; Fischer 1982; Marsden 1988; Ruef, Aldrich and Carter 2003) and in a diverse set of empirical contexts (e.g., South et al. 1982; Shrum, Cheek and Hunter 1988; Reagans 2005; Marmaros and Sacerdote 2006), research has demonstrated that people associate most with others who are similar to themselves. In organizations, social homogeneity facilitates communication, predictability, and trust (Kanter 1977). Theory and empirical evidence point to two distinct mechanisms that promote homophily: choice homophily, or the preferences of actors to affiliate with similar others, and induced homophily, which results when people find themselves in situations in which they are surrounded disproportionally by others like themselves (McPherson and Smith-Lovin 1987).

In the organizational literature, scholars have observed the uneven distribution of social groups across many organizations and argued that homophily tends to subtly reinforce social stratification by providing more beneficial social capital to members of majority groups (Baron, Davis-Blake and Bielby 1986; Bielby and Baron 1986; Ibarra 1993; Ridgeway 1997; McGuire 2002; Singh, Hansen and Podolny 2010). But two significant and important theoretical gaps remain in our current understanding of homophily in organizations. First, although we know that homophily results both from the choices of individual actors and from their uneven distribution across an organization, our understanding of exactly how these two mechanisms interact in an organizational context is under-specified. Second, we know little about how the interplay of homophily, formal organizational structure and communication patterns differ between men and women in contemporary organizations.

In this paper, we argue that the opportunity structure of possible interaction partners is strongly influenced by a firm's formal organizational structure and by its geography: people who are assigned to the same business unit, the same job function or the same office building are far more likely to interact than those who are not. But within the constraints of the organizationally and geographically determined opportunity structure, actors often have plenty of discretion to exercise choice homophily. We argue that homophily results from discretionary choice within the boundaries of the firm's opportunity structure. In effect, the organization bridges the mechanisms of homophily by establishing the opportunity sets within which choices are exercised. Consistent with this hypothesis, our results show that same-gender interactions occur at a significantly higher rate in within-business unit and within-office dyads than in between-business unit and between-office dyads, respectively. We find no support for the hypothesis that job function plays a similarly constraining role.

As an extension to our theory, we explore the existence of gender differences in response to organizational constraint. We find that the networks of men exhibit patterns consistent with our theory of discretion within the constraints of business unit and office boundaries. In contrast, women's homophilous ties occur at a higher overall rate, but are insensitive to organizational boundaries. It appears that more than men, women reach out across boundaries to interact with other women.

These findings have implications for research on homophily, gender in organizations and organization design. First, research has long distinguished between opportunity and choice as alternate mechanisms of homophily (McPherson and Smith-Lovin 1987), but the parameters of opportunity and constraint in organizations remain under-specified. We fill this gap by demonstrating that it is business unit and office boundaries that most strongly influence the opportunity set of potential interaction partners for organizational actors. Second, we contribute to the literature on gender in organizations by offering more detailed – and more current – evidence concerning gender differences in network structures. Our findings raise questions for future research about whether conventional wisdoms regarding gender differences in social network structure remain accurate in current-day organizations. Third, we contribute to the literature on organization design, offering evidence of how and by whom formal lateral structures serve to link together an otherwise siloed organization. Our finding that, in forming homophilous ties, women tend to bridge formal structural boundaries, links these heretofore disparate literatures by suggesting a new perspective on the potentially valuable role of women in knitting together the fabric of the multidivisional corporation. In conclusion, we join a rising chorus of scholars calling for greater attention to the interplay between formal and informal structure in creating organizational outcomes (e.g., Gulati and Puranam 2009; Yakubovich and Shekshnia 2009; Soda and Zaheer 2010; Tzabbar et al. 2010), but offer a unique contribution to the discussion by looking at homophily, an antecedent – not a consequence – of such interactions.

HOMOPHILY: CHOICE AND OPPORTUNITY STRUCTURE

In a broad range of relationship types and across a diverse set of contexts, researchers have demonstrated that people tend to associate more with others who are similar to themselves than with those who are not. Why should this be? One possibility is that social actors deliberately choose to associate with others who are like themselves. There is evidence of such choice homophily in friendship networks among children (Shrum, Cheek and Hunter 1988), college students (Marmaros and Sacerdote 2006) and adults (Lazarsfeld and Merton 1954); confiding networks among adults (Marsden 1988); social support networks among government bureaucrats (South et al. 1982); interaction networks among co-religionists (Fischer 1982); and work collaboration networks among entrepreneurs (Ruef, Aldrich and Carter 2003), to cite a few among myriad examples (cf. Ingram and Morris 2007). Throughout this literature, gender is one of the most salient categories in driving homophilous interaction (McPherson and Smith-Lovin 1987).

But preference is not the only reason why we observe heightened interaction rates between members of the same social categories. Similar people often sort (or are sorted) into similar situations and, as a result, people of a kind often find themselves in positions, such as jobs (Bielby and Baron 1986; Kaufman 2010), neighborhoods (Laumann 1966), or voluntary organizations (McPherson and Smith-Lovin 1987), that are disproportionally populated with others like themselves. In such an environment, even if people's choices of relationship partners were random with respect to membership in social categories, we would nevertheless observe elevated interaction rates among demographically similar people as a result of the uneven distribution of social categories across positions. When positions are relatively homogeneous and serve to focus social relations, they produce structurally induced homophily (Feld 1981; McPherson and Smith-Lovin 1987).

While the notion that homophily is induced by the opportunity structure of potential interaction partners is not novel, our contribution lies in teasing apart opportunity from choice in the intra-organizational context. Most of the relevant literature has examined homophily in society, free from the task requirements of a formal organization. For example, Bossard's (1932) classic study examined the propinquity effect on spouse selection in Philadelphia. More recently, Aral et al. (2009) examined the role of homophily in driving co-adoption of technology services by friends. While the many studies in this broad literature provide useful insights into the mechanisms and consequences of homophilous interaction, we have little assurance that they accurately illuminate the sources of homophily in the context of formal organizations, where the task structure more tightly prescribes patterns of interactions.

There has been a great deal of influential research on homophily in organizational settings, but little of it has specifically focused on distinguishing the relative contribution of opportunity versus preference as the mechanism most responsible for its emergence (cf. theoretical work by Ibarra 1993; Ibarra and Smith-Lovin 1997). Illustrative, for instance, is Blau's seminal work, in which he discusses "the internal structure of organizations, conceptualized as the distribution of their employees among official positions along various lines," (Blau 1994:130) but in which he offers no specificity about what positions or lines serve to structure interaction opportunities. Indeed, although we may have strong intuitions about how opportunity and choice come together in an organizational context, we have few empirical results on which to rely. In consequence, this work tends not to consider the potential influence of individual preference and how it may operate in conjunction with organizational mandates to determine the level of homophilous interaction. In short, we lack an integrated theory of homophily in organizations that accounts for both opportunity structure and choice in determining tie formation.

OPPORTUNITY, CONSTRAINT AND DISCRETION IN TASK PERFORMANCE

For individual preference to be a significant source of homophily in communication relationships within organizations, actors must have a certain measure of latitude to choose their interaction partners. To what extent do they in current day organizations? And, across what dimensions of organizational structures are individuals more likely to have discretion to choose their contacts, versus having them dictated by the responsibility to complete a specific set of tasks? We rely on several literatures to describe the conditions under which actors are more likely to have the freedom to choose their communication partners; in so doing, we lay out boundary conditions for our theory.

The earliest research on this question dates back to the turn of the 20th century, most notably in the work of Frederick Taylor (1911). Taylor and other adherents to "scientific management" drew a distinction between workers and managers and to maximize efficiency, they sought to reduce to zero the amount of discretion exercised by workers in performing their jobs. Subsequent perspectives acknowledge that even industrial work contains some elements that are prescribed and others that are discretionary (Turner and Lawrence 1965). More recent scholars go further, emphasizing that giving workers a greater level of discretion, even in work that is as highly scripted as a factory assembly line, can increase efficiency (Adler, Goldoftas and Levine 1999). Indeed, since McGregor (1960), there has been a long literature in human resource management showing that autonomy, which is defined as "the amount of discretion the worker is expected to exercise in carrying out assigned work activities" (Turner and Lawrence 1965:21), not only exists, but it is associated with such positive outcomes as motivation, job satisfaction and individual performance (Hackman and Oldham 1980). And as we travel across the continuum from worker to manager and from the factory floor to knowledge work, we expect to see that actors should have greater discretion in choosing with whom to communicate. Indeed, knowledge work is fundamentally dependent on the sharing of information and know-how that is distributed across an organization (Kogut and Zander 1992). As such, knowledge workers must use discretion to locate and acquire the information needed to do their jobs (Hansen 1999; Tsai 2001).

But within any organizational context, discretion in selecting communication partners is constrained. Indeed, a significant gap in the literature is the under-specification of the role of organizational structure in constraining individual discretion in interaction patterns and, in doing so, in affecting the structure of homophily. The task structure of the firm, embodied in its formal organizational structure and geographic configuration (Galbraith 1973; Tushman and Nadler 1978), form a large component of the opportunity structure within which organizational members exercise discretion to engage in interpersonal interactions. Employees often must interact with specific others to complete the task requirements of their jobs, but the degree to which this is true is likely to depend on the nature of the task and the costs of interaction. Indeed, Allen, building on Thompson (1967), argues that this assertion is true by design. He writes: "The real goal of formal organization is the structuring of communication patterns," (Allen 1977:211). In the typical, modern, complex organization there are three major types of structures that we believe are likely to sharply delimit communication patterns: strategic business units, functional units and geographic units.[1] We discuss each, in turn, to consider their role in structuring homophilous interactions.

Business Units. Chandler (1962) famously characterized many of the large organizations since the turn of the last century as adopting M-forms, in which operational decisions occur within business units and strategic decisions are managed at the headquarters level. In this view, individuals whose task requirements necessitate reciprocal interdependence are organizationally co-located within a task-oriented business unit to focus and prioritize their interactions with each other and to minimize the costs of coordination within the organization (Thompson 1967; Galbraith 1973). The business units of a multi-divisional firm are designed to be largely autonomous of one another, with interactions focused within, rather than between, them (Galbraith 1973; Williamson 1975). Thus, as a general rule, interactions between individuals in different business units likely are episodic and pertain to specific activities relative to communications among members of the same strategic business unit (SBU). This implies that interactions across business unit boundaries, *ceteris paribus*, will connect individuals in specific roles to accomplish specific tasks. Because of the specificity of the interaction, we anticipate that employees are less likely to exercise personal preferences in choosing the alters with whom in they communicate in cross-SBU interactions.

[1] Of course, SBU, function, and geography are the major categories that sort people and tasks within organizations, but potentially there are myriad, lower-level sub-divisions and structures that further delimit interaction patterns. For example, work groups, committees, task forces, and formal reporting structures are a few of the many sub-structures that will mold interactions within the supra-structures of business unit, function, and geographic collocation.

Functional Units. Within multidivisional firms, business units are further subdivided along functional lines (Galbraith 1973; Hrebiniak and Joyce 1984). These functional units serve two distinct purposes: lateral linking and a more focused division of labor. In the multidivisional firm, functional units occasionally provide a locus for interaction across business unit boundaries. For example, some organizations promote cross-SBU, within-function sharing of best practices (Galbraith 1994). However, the primary purpose of functional units is to create further specialization within each business unit, narrowing the range of tasks performed by each person, and co-locating reciprocally interdependent tasks within organizational units that are smaller than the overall SBU. In the process, the existence of functional units also sharpens the set of relevant interaction partners for each person. Thus, like business units, job functions prioritize interactions within their boundaries, relative to cross-functional interactions.

Job functions, however, differ from business units in one important respect concerning their potential influence on the incidence of within- versus across-organizational unit interaction. Like business units, job functions are a structural means to achieve a division of labor. However, unlike business units, job functions are designed to be interdependent (Williamson 1975). In a typical organization, it would be reasonable to expect a higher level of cross-function than cross-business unit interaction. Nevertheless, our belief is that most inter-functional interactions are relatively formalized, with coordination requirements tightly prescribed by the design of the organization. Thus, in spite of the fundamental interdependence that exists between job functions, we expect that, as in business units, job functions will serve to delimit individual interaction opportunities. We anticipate a relatively greater set of substitutable interaction partners within- relative to cross-functional boundaries, and therefore, as with business units, we expect that individuals will have greater discretion in selecting communication partners in intra- relative to inter-functional communication.

Geographic Units. There is evidence of significant spatial effects on the rate at which actors associate. Studies find that the probability of a relationship increases sharply when two individuals live or work near to one another (e.g., Zipf 1949; Festinger, Schachter and Back 1950; Blau and Schwartz 1984; Kono et al. 1998; Sorenson and Stuart 2001). This is true of geographic space, of functional spaces within physical structures (Festinger, Schachter and Back 1950), and of micro spaces within buildings (Marmaros and Sacerdote 2006; Liu 2010). In fact, despite rampant speculation that the proliferation of electronic communication will herald "the death of distance" (Cairncross 2001), the evidence on the issue contradicts the view that modern communication technologies have dramatically reduced the impact of geographic proximity on the likelihood of interaction (Marmaros and Sacerdote 2006; Mok, Carrasco and Wellman 2010). Whereas business units and functional structures focus interaction by proscribing specific job tasks, geographic proximity creates interaction opportunities through the local availability of convenient, low cost interaction partners. Zipf referred to mechanism behind the proximity effect as "the principle of least effort" (Zipf 1949), which underscores that the lowest cost interactions tend to be among co-located individuals.

Geographic collocation is a residual category of social organization in formal organizational contexts. It may coincide with business unit and functional memberships, as organizations often choose to geographically group individuals who share common structural units. Thus, in many organizations, the geographic distribution of individuals partly is dictated by organization structure. After accounting for affiliations to particular organizational units, however, geography captures the ease of interaction. Conditional on accounting for common organizational affiliations, we expect that collocated individuals will have a high degree of discretion in selecting local interaction partners. Indeed, there is reason to expect that the highest level of discretion in selecting communication partners will occur within office boundaries. This is because of the nature of within-office ties: relative to the other interactions within organizations, we surmise that intra-office interactions are more likely to be informal and social, including casual conversations, lunch breaks, and the like. Or, in Allen's (1977) words, intra-office interactions are more likely to be "neutral social relations".

When actors communicate across business units, job functions and offices, we suspect that they are more likely to be of an episodic nature and driven by a specific set of task requirements. This suggests that individuals will be less likely to know a broad group of potential informants for the task at hand and their interactions are more likely to be prescribed by formal task responsibility, which narrows their interactions to alters in specific job roles. Given that their choice set may be limited to those relatively few people whom they happen to know, they have less opportunity to choose their interaction partners. We therefore expect to observe less choice homophily across these boundaries then we will witness within them. We hypothesize:

> **Hypothesis 1**: Same-gender interactions will occur more frequently within <u>business units</u> than between them.
> **Hypothesis 2**: Same-gender interactions will occur more frequently within <u>job functions</u> than between them.
> **Hypothesis 3**: Same-gender interactions will occur more frequently within <u>office locations</u> than between them.

Data and Methods

SAMPLE AND DATA COLLECTION

Data for this study was collected from BigCo, a large information technology and electronics company. BigCo has 29 product divisions, organized into four primary product groups: hardware, software, technology services and business services. Overlaid across the business unit structure is a formal lateral organization, in which each person is also assigned to a job function. Within this formal structure, the employees of BigCo are widely dispersed geographically, with a relatively imprecise coupling between formal structure and geography.[2]

The data we analyze include the complete record, as drawn from the firm's servers, of e-mail communications among 30,328 employees. These data are well suited to test the hypotheses. First, because we can collect electronic communication data for large numbers of individuals at low cost, we can explore the determinants of homophilous interaction in a larger, more complex organization than those studied previously. Given our interest in the influence of organizational structure on shaping interaction, we prefer to study a multi-unit, multi-function, multi-office organization. Second, by measuring actual communication, rather than self-reports of friendship, social or instrumental ties, we are

[2] Although BigCo is global in scope, privacy laws in Europe and Asia limited our data collection to the 289 offices spread throughout the United States.

able to observe homophilous interactions directly, rather than filtering them through the subjective perceptions of survey respondents (Bernard, Killworth and Sailer 1981; Quintane and Kleinbaum 2011).

All internal e-mail information that was on the server at the time of data collection, spanning an observation period from October through December, 2006, was included in our sample. BigCo provided the data in the form of 30,328 text files, each representing the communication activity of a single person, which we cleaned and parsed. To protect the privacy of individual employees, BigCo stripped all messages of their content, leaving only the meta-data (e.g., sender, recipient, timestamp). We consolidated these files and expanded each multiple-recipient message to include one entry for each unique dyad. The final file contains 114 million e-mails.

We focus our analyses on e-mails that were sent to four or few recipients. In the core models, we exclude BCC recipients, mass mailings, and direct interactions with administrative assistants. Imposing these screens shrinks the data set by almost an order of magnitude to 13 million e-mails.[3] The mean employee in our sample exchanged 385 non-mass, non-BCC e-mails with 26 other members of the sample during the observation period (a median count of 3 and a mean of 10.9 messages within each communicating dyad), as well as 757 e-mails with employees in BigCo who are not included in our sample. These distributions have very long right tails: the maximum number of correspondents was 265 and the maximum number of e-mails was close to 14,000. BigCo also provided demographic and HR information about each employee, which we are able to link to the communication data through encrypted employee identifiers. The HR data include each employee's business unit, major job function, job sub-function, firm tenure, salary band, state, office location code and gender.

[3] Of the original 114 million dyadic e-mails, 31 million involved a person either outside the United States or otherwise not included in the sample, and about whom we have no demographic data; 3.5 million involved an administrative assistant; 1.2 million were BCCs (in these instances, we retain the message for To: and Cc: recipients, but do not treat the sender-to-BCC recipient as a realized tie). In addition, 64 million were mass mailings (i.e., they included more than four recipients). Mass mailings represent just 17% of total e-mails sent but they are just over 50% of pairwise exchanges based on the expansion of the message to include all sender-to-recipient ties.

Overall, BigCo's non-administrative U.S. workforce is 69.9% male and 30.1% female.[4] The four major product groups of the company are similar in their gender composition, ranging from 25% to 28% female. The corporate sales organization and corporate headquarters have higher proportions of women than other units, at 32% and 39%, respectively. There is also some gender sorting into job functions: in addition to administration, women are over-represented in finance and form a majority of employees in the communications and human resources functions. Conversely, men are over-represented in general executive management and research and development. The two largest functions within the company, sales and services, have gender distributions similar to the company as a whole. The proportion of women at BigCo decreases with increasing rank.[5]

The overall sample contains 24% of BigCo's total U.S. employee base, but differs from the population of the firm in several respects. Therefore, the possibility exists that use of the full sample could produce findings that are biased in unknowable ways relative to the true patterns of interaction in BigCo. To guard against the risk that our findings are driven by sampling issues, we exploit our large sample size and our knowledge of the firm's population of U.S.-based employees to create a stratified random sub-sample of employees. Our sub-sampling approach, which maximizes the correspondence between the sub-samples we draw and a set of population parameters, yields a sub-sample consisting of 15,116 employees. To assemble the representative sub-sample, we created a three-dimensional matrix of salary band (middle managers, 11, 12, 13, 14, everyone else), function (general executive management, marketing, sales, services, everyone else) and business unit (corporate headquarters, everyone else). For each of the 60 cells of this 6×5×2 matrix, we calculated the sampling probability that would be needed to achieve a sub-sample rate of 11.9% of the U.S. population of the firm. We chose to make our sub-sample representative of only selected groups in order to maintain a large sample size; had we made our sub-sample representative across the board, we would have diminished our sample to just 2.9% of the U.S.

[4] As in many U.S. corporations, the administrative staff is overwhelmingly female. Throughout this paper, however, we focus exclusively on non-administrative employees.
[5] To protect the privacy of BigCo and its employees, we cannot disclose the precise gender distribution in specific parts of the company.

population of the firm. Once we had these probabilities, we used a random number generator to determine whether each person in the overall sample, given her salary band, job function and business unit, would be included in the sub-sample. The analyses we will present are based on the more conservative random sub-sample, but the findings do not substantively change from the full sample or in various random draws of the sub-sample.

ESTIMATION APPROACH

After cleaning and parsing the data, we collapsed them into a single cross-section and created counts at the dyad level of the total number of $i \leftrightarrow j$ messages, where i and j index all individuals in the sample. In other words, we constructed a cross sectional dataset with counts of the number of communications within unordered pairs of individuals. We then undertook multivariate analyses to model the frequency of dyadic communication based on common group memberships and other pairwise attributes of each dyad. Even with the time axis compressed so that the data are structured as a single cross-section, the communication matrices are large and sparse. In fact, less than 0.5% of the approximately 114 million possible unordered cells in the sub-sample e-mail matrix are non-zero. Even given modern computing power, it is not expeditious to work with the full matrix.

Random sampling from the set of the 114 million dyads is one potential solution to this problem. However, this approach ignores the fact that the realized (non-zero) ties provide most of the information to identify the parameter estimates (Cosslett 1981; Imbens 1992; Lancaster and Imbens 1996). We therefore construct a "case cohort" dataset by including in our regression models all non-zero cells and a random sample of zero cells (King and Zeng 2001), which are then weighted according to their probability of being drawn into the analysis sample. We do not stratify on the sampling of zeros; we simply draw the zero cells at random.

Our dependent variable is a count of the number of e-mails exchanged within each dyad[6]. To accommodate the case cohort data structure, we use a weighted quasi-maximum likelihood (QML) Poisson model. Because the Poisson is in the linear exponential family, the coefficient estimates are consistent as long as the mean of the data is correctly specified. No assumptions about the distribution of the data are required[7] (Gouriéroux, Monfort and Trognon 1984; Wooldridge 1997; Silva and Tenreyro 2006). Thus, we estimate the likelihood that dyad-level covariates affect the frequency of interaction using models of the form:

$$\mathrm{E}\left[Y_{ij} \middle| X_{ij}\right] = \exp\left((X_{ij} + Z_{ij})\beta\right) \qquad (1)$$

where Y_{ij} is the count of e-mails exchanged (in both directions) between individuals *i* and *j*, X_{ij} is a vector of pair-level covariates, Z_{ij} a vector of control variables, and β a vector of regression coefficients.

In dyad-level models like the ones we estimate, there is a well known estimation problem: observations are likely to be non-independent. In particular, dyad models are prone to two types of non-independence that potentially can yield misleading results. First, interactions within a dyad are not independent: the number of e-mails actor *i* sends to actor *j* is dependent on the number of e-mails *i* receives from *j* (Quintane and Kleinbaum 2011). To address this problem, we analyze the total number of messages exchanged within the dyad; that is, the value Y_{ij} includes messages sent from *i* to *j* and from *j* to *i*. To eliminate the problem of reciprocal correlation, Y_{ji} does not appear in our analysis. Second, each individual in a dyad appears in multiple, other dyads, which introduces a common person effect (Kenny, Kashy and Cook 2006). That is, Y_{ij} may be correlated with Y_{ik} to the extent that unobservable attributes of

---

[6] In unreported results, we used two alternative specifications of the dependent variable. In one, the dependent variable was a binary indicator of whether or not the dyad members communicated during the observation window and logit models were estimated. In the second, zero-inflated Poisson models were used to separately estimate the probability of dyadic communication and its frequency, conditional on a non-zero value. Across all approaches, the overall pattern of results was substantively the same.

[7] Unlike the maximum likelihood Poisson, quasi-maximum likelihood estimation of the Poisson does *not* assume that the data are distributed with the mean equal to the variance of the event count. Unless the data are known to be distributed negative binomial, Poisson QML estimation is preferable because it is consistent even if the data are, in fact, distributed negative binomial. The only assumption of PQML concerns the distribution of the conditional mean of the data (Gouriéroux, Monfort and Trognon 1984; Wooldridge 1997; Silva and Tenreyro 2006). For robustness, we estimated negative binomial models as well and the findings are comparable.

person *i* affect both values. This problem should not affect the point estimates, but can cause us to underestimate the values of standard errors (Kenny, Kashy and Cook 2006).

We address the non-independence problem by estimating robust standard errors that are simultaneously clustered on *both* members of a dyad. Cameron, Gelbach and Miller (2011) develop this approach theoretically, but only implement it for ordinary least squares and logit regression. Because their approach is more generally robust, we develop a more general implementation of it in Stata that is suitable for other regression estimators, including the Poisson quasi-maximum likelihood models we employ. As in Cameron, Gelbach and Miller (2011), standard errors are calculated in three, separate, cluster-robust covariance matrices: one by clustering according to *i*, one by clustering according to *j*, and one by clustering according to their intersection. Standard errors in the regressions we report, which cluster on both dyad members, are estimated based on the matrix formed by adding the first two covariance matrices and subtracting the third. This approach is similar to using the quadratic assignment procedure to adjust standard errors in multiple regression (MR-QAP) [8], but it can be implemented more quickly in large data sets (Cameron, Gelbach and Miller 2011). Likewise, this approach is feasible on much larger data sets than can be analyzed using exponential random graph modeling (ERGM).

INDEPENDENT VARIABLES

The independent variables in our dyad-level regressions are all properties of the *ij*th pair of employees. Of primary interest in our analysis is a set of dummy variables that indicate whether or not two individuals, employees *i* and *j*, share the same affiliation across six different organizational and social groups. First, we include *SameBU*, defined to be one when *i* and *j* are in the same strategic business unit. BigCo has 29 business units, which are organized into four business groups; additionally, corporate headquarters and the corporate sales force are treated as business units by the company and in our data. We include a *SameFunction* dummy variable to indicate whether employees *i* and *j* are in the same job

[8] MR-QAP adjusts standard errors to correct for violation of the independence assumption through a re-sampling method, similar to the bootstrap (Simpson 2001).

function. BigCo classifies each employee in one of 13 different job functions: administration (consisting primarily of secretaries and other support personnel), communications, finance, general executive management, human resources, legal, manufacturing, marketing, research & development, supply chain, sales, services and a catch-all "other" category. These 13 job functions are further sub-divided into 60 subfunctions, which we account for in the regressions with a *SameSubfunction* dummy variable.

Employees in our sample work in 289 offices scattered across all 50 U.S. states. We include a *SameOffice* dummy to indicate pairs of actors who are physically located in the same building. We also include *logDistance*, the natural logarithm of the estimated door-to-door (driving) distance between employee *i* and employee *j*'s office, plus one mile. The company has a 15-band salary hierarchy ranging from 0 (for employees in training) to 14. We include a *SameBand* dummy variable to indicate that both members of a dyad are in the same salary band. We have two different dummy variables for gender; in our first set of models, we include *SameGender*, set to 1 when *i* and *j* are either both male or both female, and 0 otherwise. In our second set of models, we split out the same gender effect by including separate covariates for *BothMale* and *BothFemale*[9]. For confidentiality reasons, BigCo would not disclose other socio-demographic data, such as employees' race, ethnicity or age.

Our hypotheses anticipate that homophily will be stronger within geographic and formal organizational units than between them. To test these predictions, we add to our baseline model interaction terms in successive models: *SameGender*×*SameBU, SameGender*×*SameOffice*, and *SameGender*×*SameFunction*. Because both variables in each interaction term are dummy variables, the interaction term is tantamount to a difference-in-differences estimator (Ashenfelter and Card 1985). The estimator tests the null hypothesis that the difference in communication frequency between same-gender and different-gender dyads is equal within compared to across units. Thus, a positive, significant

[9] We use this contrast coding system (Kaufman and Sweet 1974) to highlight differences between both-male and male-female dyads and between both-female and male-female dyads, respectively; in unreported results, we replicate our analyses re-specifying our dummy variables as *SameGender* and *BothFemale* to get explicit significance tests for differences between both-male and both-female dyads.

interaction term would be sufficient to reject the null hypothesis and conclude that gender homophily is stronger within geographic and formal organizational units than between them.

In addition to examining gender homophily overall, we also look separately at homophily between men and homophily between women. To do this, we drop the covariate *SameGender* and replace it with covariates for the main effects of *BothMale* and *BothFemale* as well as the interaction variables *BothMale*×*SameBU* and *BothFemale*×*SameBU*. The former can be considered as a difference-in-differences estimator to test the hypothesis that the difference in communication frequency between male-male dyads and mixed-gender dyads is greater within business units than across business units. Conversely, the latter interaction variable tests the hypothesis that the difference in communication frequency between female-female dyads and mixed-gender dyads is greater within business units than across business units.

CONTROL VARIABLES

We control for a set of variables we expect will affect the propensity for dyadic communication. Once again, because all regressions are performed at the dyadic level of analysis, the covariates all are specified at the level of the pair rather than the individual. For all categories in the regressions, we control for the combined sizes of the groups to which the members of the dyad belong. These group sizes define the risk set of possible local and cross-group communication partners. For instance, we include *logprodBUSize*, the natural logarithm of the product of the number of people in the sample who are in the business units to which i and j belong[10]. In general, when i and j are members of large business units (or other large groups), the probability that they specifically will interact will decline because of the large number of available substitute communication partners (assuming that individuals' interaction frequencies do not scale proportionately with group size). We include similar group size controls for function,

[10] In using this specification, we build on a robust econometric literature on gravity models of world trade, whose inspiration comes from models in geophysics (Stewart 1941). Econometricians have long debated the correct specification of mass in these gravity models (analogous to group size in our models); the most current consensus suggests that the log of the product of the group sizes is the correct specification (Carrère 2006). However, our results are robust to other specifications, such as the log of the average (or cumulative) group sizes.

subfunction, and office. Finally, we control for the gender distribution of the environment surrounding the dyad by including, for example, *BUPctWomen*, the cumulative percentage of females in *i*'s and *j*'s business units. We include similar controls for the gender distribution of *i*'s and *j*'s home office and job subfunction. In addition to clustering by actor, we also control for communication volumes directly.[11] We include *logEmailVolume*, the natural logarithm of one plus the number of e-mails the two actors exchanged with all other (non-*i*-*j*) partners in our sample to adjust for the fact that the individuals within the sample have differential propensities to communicate.

Results and Discussion

Descriptive statistics and a correlation matrix of dyad-level variables are shown in Table 1. Table 2 shows descriptive statistics of individual-level e-mail activity, broken out by gender. Surprisingly, we find that women, on average, have higher total e-mail volume than men. The average woman at the company was involved in 2,850 messages during the observation period, compared to 2,564 for the average man. This higher total results both from a larger number of contacts and a larger average frequency of interaction within each dyad[12]. We also find that women, on average, have a higher proportion of female contacts than do men. Table 2 shows striking gender differences in patterns of homophilous interaction. The average male in BigCo maintains an overall contact distribution that reflects the gender composition of the company almost precisely: 70% of his contacts are other men, and 30% are women. Conversely, women are over-represented in the communication networks of women: 42.2% of a typical female's contact network is female, whereas only 30% of BigCo employees are female. Note that

---

[11] If we exclude these volume controls, our results would indicate additional communications in which men or women engage on the margin, as opposed to shifts in the distribution of a fixed number of communications across potential recipients. (As we report below, there are, in fact, significant gender differences in the volume of communication: women at BigCo send and receive more e-mails than men do.) Unreported results reveal the same general pattern of results in models that exclude communication volume controls. While we take comfort in the similar findings, we prefer to include volume controls to better account for unobserved heterogeneity in communication behavior.

[12] We also note that these differences are even larger when we adjust for rank because both e-mail volume and number of contacts tend to increase with salary band and women, in aggregate, have a lower average salary band than do men.

this difference is large and it exists because women have considerably more contacts then men. The additional contacts that women initiate are far more likely to be with other women.

Table 2 also reports the gender composition of interactions within versus across organizational units. The numbers in this table are purely descriptive; we do not adjust the cells for any factors. Results indicate that for the average man in the sample, the gender distribution of his contacts is shifted slightly in favor of men for interactions within his own business unit (70.6%), office location (72.0%) and job function (71.7%), relative to the overall distribution of 69.9% male. Conversely, the average man's interactions across business units, offices, and functions include a slightly lower proportion of same-gender ties. Similarly, the average woman interacts with other women in her own business unit at a slightly higher rate, compared to across business units. But when we examine office and function boundaries, the pattern of results reverses. We are surprised to observe that women have a higher proportion of their cross-office (42.3%) and cross-functional (44.8%) ties with women, relative to their within-office (41.2%) and within-function (42.9%) contacts.

Tables 3 and 4 present dyad-level Poisson quasi-maximum likelihood regression models of the frequency of e-mail exchange in BigCo. Models 1-4 in Table 3 consider gender homophily generally; Models 5-8 in Table 4 differentiate between male and female gender homophily. Models 1 and 5 provide baseline estimates for the two approaches, respectively, and both show positive, significant (both statistically and substantively) main effects on communication of co-membership in formal structural groups (business units and job functions), geographic units (offices) and social categories (gender). The single largest organizational effect on the rate of communication is sharing the same business unit affiliation. When individuals i and j are in the same business unit, they interact at exp(2.352)=10.51 times the rate of otherwise similar dyads that span different business units. The effects of being in the same function and subfunction are large as well: two individuals in the same function communicate at exp(1.019)=2.77 times the rate of those who are in different functions, *ceteris paribus*. Two individuals who also are in the same subfunction communicate at 3.31 [= exp(1.198)] times the rate of those who are in the same function but not the same subfunction; combining these effects indicates that those in the same function and subfunction communicate at 9.17 [= exp(1.019 + 1.198)] times the baseline rate.

Turning next to geography, we include a dummy variable indicating whether employees *i* and *j* are in the same office and the log of the distance between them. *SameOffice* has a very large effect on the rate: two individuals communicate at [exp(1.032 × *SameOffice* – 0.186×ln(distance + 1 mile))] times the rate as otherwise identical, cross-office pairs. Compared to two employees separated by just 100 miles, two people in the same office communicate 6.62 times more frequently. Relative to a dyad separated by the mean geographic distance in the sample (1,026 miles), two people in the same office communicate at 10.19 times the rate. Finally, we consider sociodemographic categories. The gender composition of the population (and of our subsample) is 70.1% male and 29.9% female. We do find a positive effect of gender homophily: overall, same-gender dyads exchange e-mails at a rate 17% higher [1.17 = exp(0.155)] than that of mixed-gender pairs. When we split this effect by gender, we find that male-male dyads exchange e-mails at a rate 14% higher [1.14 = exp(0.129)] than that of mixed-gender pairs, whereas in female-female dyads, the corresponding number is 26% [1.26 = exp(0.230)].

We control for the gender distribution of the task environments of the dyad members. Our *OfficePctWomen* and *BUPctWomen* controls were not significant, but at a more local level, there was a positive, significant effect – albeit modest in magnitude – of *SubfunctionPctWomen* in all models. For every one percentage point increase in the cumulative proportion of women in the dyad members' job subfunctions, there was a 0.9% increase in dyadic communication frequency [1.009 = exp(0.009)]. This could be an artifact of the higher overall communication volume of women and – because we control for the gender of dyad members in Models 5-8 – might indicate a contagion in interaction in groups that have relatively many women.

HYPOTHESIS TESTS

To test our hypotheses that actors are more likely to choose same-gender communication partners within the opportunity structure that is defined by business units, job functions and office buildings, we introduce to the baseline models interactions between co-membership in formal structural units and co-membership in social categories. In Models 2 and 3, we find positive, significant coefficients on the interactions of *SameGender* with *SameOffice* and *SameBU*, respectively. These interactions are displayed graphically in the figures in Table 3. We find that relative to same-gender dyads who are <u>not</u> in the same office, same-gender dyads who <u>are</u> in the same office communicate 3.0 times as much [= exp($\beta_{SameOffice}$ + $\beta_{SameGender \times SameOffice}$) = exp(0.904 + 0.209)]. This is ratio is significantly higher than that implied by the main effect of *SameOffice* alone [2.5 = exp(0.904)]. Likewise, relative to same-gender dyads who are <u>not</u> in the same business unit, same-gender dyads who <u>are</u> in the same business unit communicate 11.0 times as much [= exp(β_{SameBU} + $\beta_{SameGender \times SameBU}$) = exp(2.285 + 0.114)]. This is ratio is significantly higher than that implied by the main effect of *SameBU* alone [9.8 = exp(2.285)]. Co-membership in office or business unit each amplifies the effect of gender homophily. Thus, consistent with Hypotheses 1 and 3, office and business unit boundaries provide the opportunity structure for communication; within the constraints of that structure, individuals appear to exercise more choice homophily. Model 4 in Table 3 examines whether job function plays a similar role in structuring communication. Contrary to Hypothesis 2, we find no significant *SameGender*×*SameFunction* interaction in Table 3. Therefore, Hypothesis 2 is rejected.

Why might the results for job function differ from those for business unit? One possibility is that, unlike business units that typically silo interactions within their boundaries and limit cross-unit interactions, and unlike shared geographic spaces that facilitate local, discretionary social interactions through the ready availability of communication partners, functional boundaries afford a relatively equal amount of discretion in the choice of interaction partners both locally and remotely. As we rethink the argument in light of the findings, the organizational design literature suggests that a key distinction between a functional structure and a divisional structure is that job functions have explicit interdependence requirements between them (Galbraith 1973; Hrebiniak and Joyce 1984). In contrast, divisional units typically have no such interdependence requirements and, indeed, are argued to be better left independent of one another (Williamson 1975). We expected that within a large business unit of a large organization, interactions would be concentrated within the business unit and further concentrated within the job function, with tightly prescribed inter-functional interactions that leave little latitude in the choice of interaction partners. Because individuals will have thicker networks within than across these boundaries, they will have greater freedom to choose interaction partners based on interpersonal similarities when they are communicating within organizational units. But given the interdependence inherent in the functional structure, it appears that individuals do maintain discretionary contacts in other functions.

EXTENSION: MALE-FEMALE DIFFERENCES

To tease apart gender homophily among males from gender homophily among females, we turn to Table 4, in which we replace the *SameGender* variable with two variables: *BothMale* and *BothFemale*, each of which is estimated relative to the baseline of a mixed-gender dyad. In Models 6 and 7, we examine both their main effects and their interactions with either *SameOffice* or *SameBU*, respectively. Our results in male-male dyads echo the overall same-gender effect: we find positive, significant effects of the *SameOffice*×*BothMale* and *SameBU*×*BothMale* interactions on dyadic communication. Relative to male-male dyads who are <u>not</u> in the same office, male-male dyads who are <u>are</u> in the same office communicate 3.3 times as much [= $\exp(\beta_{SameOffice} + \beta_{BothMale \times SameOffice}) = \exp(0.898 + 0.300)$]. Relative to male-male dyads who are <u>not</u> in the same business unit, male-male dyads who <u>are</u> in the same business unit communicate 11.6 times [=exp(2.286 + 0.162)] as much. These ratios are significantly higher than those implied by the main effect of *SameOffice* [2.45 = exp(0.898)] or *SameBU* [9.84 = exp(2.286)], respectively. Thus, consistent with the theory and with our empirical results on the gender co-mingled interaction effects, men seem to exhibit more homophilous interactions in their within-office and within-business unit communications than in their cross-office and cross-business unit communications.

When we examine women's networks, however, the picture changes. Whereas the *BothMale* communication effect, relative to mixed-gender dyads, is amplified within offices and business units, we see no such interaction among women. The *SameBU*×*BothFemale* and *SameOffice*×*BothFemale* interactions are both indistinguishable from zero ($p > 0.50$). Men exhibit stronger homophily within the choice sets of potential interaction partners in their own office building or business unit relative to across these two boundaries, but the tendency of women to interact with one another is equally strong within and across office or business unit boundaries. As in the gender co-mingled results in Table 3, we find no significant interaction of *SameFunction* with either *BothMale* or *BothFemale* (Table 4, Model 8).

These relationships are illustrated graphically in the Table 4 Figures. The top panel shows how the *SameOffice* effect varies in mixed-gender, male-male and female-female dyads, *ceteris paribus*. The baseline rate of communication in mixed-gender dyads is shown in triangle markers connected by a dashed line: the rate of communication is 2.44 times higher when a mixed-gender dyad is co-located in the same office. Male-male dyads are shown as X's connected by a thin line; whereas the rate of male-male communication in cross-office dyads is elevated by just 6.7% [exp(0.065) = 1.067], the rate of male-male communication in same-office dyads is 3.53 times as high [$= \exp(\beta_{BothMale} + \beta_{SameOffice} + \beta_{BothMale\times SameOffice}) = \exp(0.065 + 0.898 + 0.300)$]. The chart shows clearly that the lines are not parallel, indicating a significant ($p < 0.01$) interaction effect. Female-female dyads are shown as solid circles connected by a thick line. Although the base rate of female homophily is large, there is no significant interaction with same office; the thick pink line is parallel to the dashed line.[13]

Discussion and Conclusion

We have long known that homophily is a twin-engine phenomenon. The motors are individual preferences to interact with similar others and differential opportunities to associate based on how people sort into physical and social locations. We argue that in an organizational context, the determinants of

[13] Because results show that the effects of the interactions with *BothMale* are significantly larger than zero, but the effects of the interactions with *BothFemale* are negative (and insignificant), logic suggests that the effect is stronger for men than for women. For confirmation, we employed a contrast coding scheme (Kaufman and Sweet 1974) that makes this comparison in explicit statistical terms, by including two dummies for *SameGender* and *BothFemale*; unreported results indicate that the *SameGender* dummy, which is estimated on variance provided by both male-male and female-female dyads, is positive and significant, and the *BothFemale* dummy, which explicitly indicates the difference between male-male and female-female dyads, is negative and significant. The effect size for female-female dyads, calculated by adding the *SameGender* and *BothFemale* effects, is indistinguishable from zero, exactly as in the results shown here. The two coding schemes are mathematically equivalent; we show these results because they are more intuitive to understand.

opportunity have been under-specified in the literature. To fill this gap, we assert that opportunity is delimited by the business unit structure and geographic dispersion of the organization. Given this opportunity structure, we argue that individuals have more discretion and therefore engage in more homophily within those boundaries than across them. Our empirical analysis of e-mail communications among employees of a large information technology firm lends support to this theory, but not to the hypothesis of a similar effect of job functions on structuring interactions.

At a more nuanced level, however, we explored male-female differences in communication patterns, and the most surprising findings in the paper concern these differences. First, in BigCo's email network, women have greater numbers of contacts than men and communicate with their partners more often than do men. We cannot know for sure what to make of this finding. It may indicate systematic gender differences in the use of e-mail (relative to other communication media), but we found no suggestion of this in our exploratory interviews at BigCo. Alternatively, it could be an indication of substantive differences in communication style or network structure. Our data do not allow us to adjudicate between these explanations. Second, there are significant differences in how gender interacts with organizational and geographic boundaries to influence the level of homophily in communications. Whereas men, consistent with our theory, exhibit more homophilous communication within business unit and office boundaries than across them, women are equally homophilous within and across those boundaries (i.e., there is no significant interaction between *BothFemale* and either *SameOffice* or *SameBU*).

There are at least two possible, albeit very different, explanations for these results. The first is an inherent gender differences argument, in which women have networks that are more collaborative and less focused on parochial interests or within-unit loyalties than are men (cf. Borgatti and Cross 2003). Recent research suggests that culturally influenced implicit attitudes toward collaboration may indeed affect the structure of social networks (Srivastava and Banaji 2011), and there may be a systematic gender difference in implicit attitudes. In fact, there is work to suggest that women tend to be more collaborative than men (e.g., Eagly and Carli 2003). Bringing these literatures together, the possibility exists that there is an unobservable propensity to collaborate that is correlated with gender and that leads women to transcend the organizational boundaries that constrain men.

A second possible explanation for this result is that women communicate with one another more often because they are effectively excluded from the male power structure. The classic literature on gender and managerial networks argued that women are largely excluded from the "shadow structure … of power" in organizations (Kanter 1977: 164). Although the situation women encounter in the labor market has clearly improved in the years since Kanter's analysis of "Indsco," (Kaufman 2010) such arguments are not limited to the annals of business history. In a more recent incarnation, Groysberg (2010) argues that women face "institutional barriers" to creating strong working relations. If women respond to all-male cliques by creating ties with other women (Pfeffer and Salancik 1978; Casciaro and Piskorski 2005), even those outside their business units, job functions and offices, that too may explain our results.

We cannot adjudicate the extent to which each of these explanations is driving our results. Regardless, we are struck by the finding that whereas men engage in homophilous interaction within the opportunity structure created by organizational structure, women are prone to connect with other women who are <u>outside</u> their business units and offices. This result builds on – and brings new empirical support to – earlier literatures. For example, Alderfer's theory of inter-group relations (Alderfer and Smith 1982; Alderfer 1987) suggests that people are simultaneously part of intersecting identity groups, defined, for example, by gender or race, and organizational groups, such as divisions or offices. When such groups intersect, people tend to use common group affiliation along one dimension to connect with people who differ from themselves along another dimension. Blau and Schwartz (1984) similarly argue that cross-cutting social circles (Simmel 1902) promote formation of boundary-spanning ties. To the extent that gender identity is more salient among women than among men, it would be expected to drive gender-homophilous relations across formal organizational boundaries among women at a higher rate than among men (Blau 1979).

This finding is important because it calls into question one of the conventional wisdoms, which is that homophily among marginalized actors tends to reinforce the stratification of power in organizations (Brass 1985). In a viewpoint that has become largely representative of the field, Ibarra argues: "homophily may serve to decrease women's access to organization-wide networks," (Ibarra 1992: 425).[14] We find, however, that women are both more central and more broadly connected in the internal BigCo email network. In fact, the data show that the average female employee of BigCo has *almost exactly* the number of men in her network as does the average male employee: women have 26.9 male contacts on average, while men have 26.7 male contacts. But the typical female employee also has an average of 19.6 women in her network, whereas the typical male has only 11.5 women in his. These findings are consistent with the possibility that lateral communication among women serves to reinforce, not undermine, their positions in the organization. Thus, our results suggest – though we do not claim that they prove – that homophilous interaction can actually help to span formal organizational boundaries that are otherwise difficult to traverse. Consequently, a more complete picture of both formal and informal structure reveals a situation in which homophily with respect to one category (in our study, gender) can actually serve to broaden a person's network with respect to another category (office and business unit)[15].

This contribution raises the question of whether, or under what conditions, boundary spanning ties between women indeed yield benefits to either the women themselves or to the organization. Prior research suggests that those who span boundaries – and more generally, people who connect otherwise disconnected others – play an important role in facilitating the flow of information within organizations, benefiting both the organization (Tushman 1977) and themselves (Burt 1992). This raises the puzzle of

[14] In a recent article, Singh, Hansen and Podolny (2010) concluded that "the world is not small for everyone." They argue that even as the small world structure of organizational social networks implies short average path lengths, women (as well as new employees others who are peripheral to the network) are less able to locate experts within the organization because they tend to begin their searches by contacting others who, like themselves, are marginalized. However, their empirical approach – a "small world" experiment in expertise location – requires each person to choose exactly one contact. Our results suggest that their findings might differ if the full network within which individuals are embedded is taken into account.

[15] Although our data are limited to gender homophily, our finding echoes similar results by Thomas (1990) and by Ibarra (1995), who showed that homophily with respect to race led minority managers to have more contacts outside their departments or groups than white managers in a purposeful effort to establish relationships with others of their own race.

why women may be less likely to convert boundary spanning ties into either personal or organizational benefit, even as they span organizational boundaries in their ties to other women, connecting otherwise disconnected populations, at a higher rate than do men. More research in this area is needed to understand how gender differences in the pattern of social interaction affect the relationship between social capital and both individual and organizational outcomes.

We conclude with several notes of caution. First, and most significantly, our data describe just a single firm. Although we analyze a vast dataset, we have no basis for any claim of generalizability beyond the single organization we study. This limitation is particularly important because we cannot know the degree to which our findings of gender differences depend on the particular composition and organizational structure of the firm we study. We do not believe that our results are an artifact of an idiosyncratic structure – indeed, a multi-divisional structure is typical of large, complex firms – but we cannot claim that our results apply to any firm but BigCo. Furthermore, as Ely (1994) has demonstrated, gender dynamics are played out on a stage in which the overall gender composition of an organization matters, in contrast to purely person-centered views in which gender roles are assumed to be intrinsic. We take comfort in the fact that BigCo's gender composition is almost surely typical of U.S.-headquartered companies in its sector, but we really cannot know whether or how the findings on the differing network structures of men and women may be influenced by gender composition in unknowable ways.

Second, although we have uniquely detailed intra-organizational data, we have only relatively coarse descriptors of formal organizational structure. We cannot refute the possibility that gender-based sorting into lower-level structures, such as work teams, may be driving some of the gender interaction patterns we observe (see, for example, Reagans, Zuckerman and McEvily 2004). Similarly, although we found strong evidence of gender differences in interaction patterns, we cannot know whether these results generalize to homophily along lines other than gender. Unfortunately, gender is the only sociodemographic variable available to us. We do believe that our conceptualization of homophily as discretionary choice within the constraints of an opportunity structure posed by formal organization and geography would extend to homophily along the lines of race, ethnicity and age cohort, but we have no empirical evidence to address this question.

Lastly, we must note that there are some drawbacks with the use of electronic mail data, although we believe that these are outweighed by compensatory factors. Most importantly, the e-mail network does not measure social relations *per se*, but communication. Neither do we know the content of the communications that occur, nor can we differentiate interactions that are purely task-related from those that are more social in nature. Indeed, we have no doubt that among the more than one hundred million e-mails in the data, a non-trivial percent discuss lunch plans, recent television shows or an office basketball pool. Therefore, we cannot classify either messages or relationships into types. Conversely, the data have real strengths. First, the sample is large; it includes millions of interactions among tens of thousands of employees. Second, it is unaffected by problems of non-response, recall or bias in survey response (Killworth and Bernard 1976; Quintane and Kleinbaum 2011). Third, for many of the questions that will interest students of organizations, we believe that the availability of what may well be the majority of interactions in an organization is as much a benefit as it is a source of concern. Indeed, communications that are purely social in nature are indications of what Allen (1977) calls neutral social interactions: even if these interactions are not themselves generating productive output for the company, they indicate to the analyst – and reaffirm to the individuals themselves – an existing interpersonal relationship that makes each person a potential candidate to help the other person meet her discretionary informational needs. Thus, as organizational theorists we believe that the myriad, incidental social interactions that occur within organizations ultimately may be as important to the productive effort of the enterprise as are purely task-driven communications.

REFERENCES

Adler, P.S., B. Goldoftas, D.I. Levine. 1999. Flexibility versus efficiency? A case study of model changeovers in the Toyota production system. *Organization Science* **10**(1) 43-68.

Alderfer, C.P. 1987. An intergroup perspective on group dynamics. J.W. Lorsch, ed. *Handbook of Organizational Behavior*. Prentice-Hall, Englewood Cliffs, NJ, 190-222.

Alderfer, C.P., K.K. Smith. 1982. Studying Intergroup Relations Embedded in Organizations. *Administrative Science Quarterly* **27**(1) 35-65.

Allen, T.J. 1977. *Managing the flow of technology : technology transfer and the dissemination of technological information within the R&D organization*. MIT Press, Cambridge, MA.

Aral, S., L. Muchnik, A. Sundararajan. 2009. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences* **106**(51) 21544-21549.

Ashenfelter, O., D. Card. 1985. Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs. *The Review of Economics and Statistics* **67**(4) 648-660.

Baron, J.N., A. Davis-Blake, W.T. Bielby. 1986. The Structure of Opportunity: How Promotion Ladders Vary Within and Among Organizations. *Administrative Science Quarterly* **31**(2) 248-273.

Bernard, H.R., P.D. Killworth, L. Sailer. 1981. Summary of Research on Informant Accuracy in Network Data and the Reverse Small World Problem. *Connections* **4**(2) 11-25.

Bielby, W.T., J.N. Baron. 1986. Men and Women at Work: Sex Segregation and Statistical Discrimination. *The American Journal of Sociology* **91**(4) 759-799.

Blau, P.M. 1979. A Fable about Social Structure. *Social Forces* **58** 777-784.

Blau, P.M. 1994. *Structural contexts of opportunities*. University of Chicago Press, Chicago.

Blau, P.M., J.E. Schwartz. 1984. *Crosscutting social circles : testing a macrostructural theory of intergroup relations*. Academic Press, Orlando.

Borgatti, S.P., R. Cross. 2003. A Relational View of Information Seeking and Learning in Social Networks. *Management Science* **49**(4) 432-445.

Bossard, J.H.S. 1932. Residential Propinquity as a Factor in Marriage Selection. *The American Journal of Sociology* **38**(2) 219-224.

Brass, D.J. 1985. Men's and Women's Networks: A Study of Interaction Patterns and Influence in an Organization. *The Academy of Management Journal* **28**(2) 327-343.

Burt, R.S. 1992. *Structural holes : The social structure of competition*. Harvard University Press, Cambridge, MA.

Cairncross, F. 2001. *The death of distance: how the communications revolution is changing our lives*, Completely new ed. Harvard Business School Press, Boston.

Cameron, A.C., J.B. Gelbach, D.L. Miller. 2011. Robust Inference with Multi-way Clustering. *Journal of Business and Economic Statistics* **29**(2) 238-249.

Carrère, C. 2006. Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model. *European Economic Review* **50**(2) 223-247.

Casciaro, T., M.J. Piskorski. 2005. Power Imbalance, Mutual Dependence, and Constraint Absorption: A Closer Look at Resource Dependence Theory. *Administrative Science Quarterly* **50**(2) 167-199.

Chandler, A.D. 1962. *Strategy and Structure: Chapters in the History of the Industrial Enterprise*. M.I.T. Press, Cambridge, MA.

Cosslett, S.R. 1981. Maximum Likelihood Estimator for Choice-Based Samples. *Econometrica* **49**(5) 1289-1316.

Eagly, A.H., L.L. Carli. 2003. The female leadership advantage: An evaluation of the evidence. *The Leadership Quarterly* **14**(6) 807-834.

Ely, R.J. 1994. The Effects of Organizational Demographics and Social Identity on Relationships among Professional Women. *Administrative Science Quarterly* **39**(2) 203-238.

Feld, S.L. 1981. The Focused Organization of Social Ties. *The American Journal of Sociology* **86**(5) 1015-1035.

Festinger, L., S. Schachter, K. Back. 1950. *Social pressures in informal groups: a study of human factors in housing*, 1st ed. Harper, New York.

Fischer, C.S. 1982. *To dwell among friends: personal networks in town and city*. University of Chicago Press, Chicago.

Galbraith, J.R. 1973. *Designing complex organizations*. Addison-Wesley, Reading, MA.

Galbraith, J.R. 1994. *Competing with flexible lateral organizations*, 2nd ed. Addison-Wesley, Reading, MA.

Gouriéroux, C., A. Monfort, A. Trognon. 1984. Pseudo Maximum Likelihood Methods: Theory. *Econometrica* **52**(3) 681-700.

Groysberg, B. 2010. *Chasing stars : the myth of talent and the portability of performance*. Princeton University Press, Princeton, N.J.

Gulati, R., P. Puranam. 2009. Renewal Through Reorganization: The Value of Inconsistencies Between Formal and Informal Organization. *Organization Science* **20**(2) 422-440.

Hackman, J.R., G.R. Oldham. 1980. *Work redesign*. Prentice Hall.

Hansen, M.T. 1999. The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits. *Administrative Science Quarterly* **44**(1) 82-111.

Hrebiniak, L.G., W.F. Joyce. 1984. *Implementing strategy*. Macmillan, New York.

Ibarra, H. 1992. Homophily and Differential Returns: Sex Differences in Network Structure and Access in an Advertising Firm. *Administrative Science Quarterly* **37**(3) 422-447.

Ibarra, H. 1993. Personal Networks of Women and Minorities in Management: A Conceptual Framework. *Academy of Management Review* **18**(1) 56-87.

Ibarra, H. 1995. Race, Opportunity, and Diversity of Social Circles in Managerial Networks. *Academy of Management Journal* **38**(3) 673-703.

Ibarra, H., L. Smith-Lovin. 1997. New directions in social network research on gender and organizational careers. C.L. Cooper, S.E. Jackson, eds. *Creating Tomorrow's Organizations: A Handbook for Future Research in Organizational Behavior*. John Wiley & Sons, New York, 359-383.

Imbens, G.W. 1992. An Efficient Method of Moments Estimator for Discrete Choice Models With Choice-Based Sampling. *Econometrica* **60**(5) 1187-1214.

Ingram, P., M.W. Morris. 2007. Do People Mix at Mixers? Structure, Homophily, and the â€œLife of the Partyâ€ . *Administrative Science Quarterly* **52**(4) 558-585.

Kanter, R.M. 1977. *Men and women of the corporation*. Basic Books, New York.

Kaufman, D., R. Sweet. 1974. Contrast Coding in Least Squares Regression Analysis. *American Educational Research Journal* **11**(4) 359-377.

Kaufman, R.L. 2010. *Race, gender, and the labor market : inequalities at work*. Lynne Rienner Publishers, Boulder, Colo.

Kenny, D.A., D.A. Kashy, W.L. Cook. 2006. *Dyadic data analysis*. Guilford Press, New York.

Killworth, P., H. Bernard. 1976. Informant Accuracy in Social Network Data. *Human Organization* **35**(3) 269-286.

King, G., L. Zeng. 2001. Logistic Regression in Rare Events Data. *Political Analysis* **9**(2) 137-163.

Kogut, B., U. Zander. 1992. Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization Science* **3**(3, Focused Issue: Management of Technology) 383-397.

Kono, C., D. Palmer, R. Friedland, M. Zafonte. 1998. Lost in Space: The Geography of Corporate Interlocking Directorates. *The American Journal of Sociology* **103**(4) 863-911.

Lancaster, T., G. Imbens. 1996. Case-control studies with contaminated controls. *Journal of Econometrics* **71**(1-2) 145-160.

Laumann, E.O. 1966. *Prestige and association in an urban community; an analysis of an urban stratification system*. Bobbs-Merrill, Indianapolis.

Lazarsfeld, P.F., R.K. Merton. 1954. Friendship as Social Process: A Substantive and Methodological Analysis. M. Berger, ed. *Freedom and Control in Modern Society*, 18-66.

Liu, C.C. 2010. Essays on network antecedents in a knowledge production context, Harvard University, Boston, MA.

Marmaros, D., B. Sacerdote. 2006. How Do Friendships Form? *Quarterly Journal of Economics* **121**(1) 79-119.

Marsden, P.V. 1988. Homogeneity in confiding relations. *Social Networks* **10**(1) 57-76.

McGregor, D. 1960. *The Human Side of the Enterprise*. McGraw-Hill.

McGuire, G.M. 2002. Gender, Race, and the Shadow Structure. *Gender & Society* **16**(3) 303-322.

McPherson, J.M., L. Smith-Lovin. 1987. Homophily in Voluntary Organizations: Status Distance and the Composition of Face-to-Face Groups. *American Sociological Review* **52**(3) 370-379.

McPherson, J.M., L. Smith-Lovin, J.M. Cook. 2001. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology* **27**(1) 415-444.

Mok, D., J.-A. Carrasco, B. Wellman. 2010. Does Distance Still Matter in the Age of the Internet? *Urban Studies* **47**(13) 2747-2783.

Pfeffer, J., G.R. Salancik. 1978. *The External Control of Organizations : A Resource Dependence Perspective*, 2003 ed. Stanford Business Classics, Stanford, CA.

Quintane, E., A.M. Kleinbaum. 2011. Matter Over Mind? E-mail Data and the Measurement of Social Networks. *Connections* **31**(1) 22-46.

Reagans, R. 2005. Preferences, Identity, and Competition: Predicting Tie Strength from Demographic Data. *Management Science* **51**(9) 1374-1383.
Reagans, R., E. Zuckerman, B. McEvily. 2004. How to Make the Team: Social Networks vs. Demography as Criteria for Designing Effective Teams. *Administrative Science Quarterly* **49**(1) 101-133.
Ridgeway, C.L. 1997. Interaction and the Conservation of Gender Inequality: Considering Employment. *American Sociological Review* **62**(2) 218-235.
Ruef, M., H.E. Aldrich, N.M. Carter. 2003. The Structure of Founding Teams: Homophily, Strong Ties, and Isolation among U.S. Entrepreneurs. *American Sociological Review* **68**(2) 195-222.
Shrum, W., N.H. Cheek, Jr., S.M. Hunter. 1988. Friendship in School: Gender and Racial Homophily. *Sociology of Education* **61**(4) 227-239.
Silva, J.M.C.S., S. Tenreyro. 2006. The Log of Gravity. *Review of Economics and Statistics* **88**(4) 641-658.
Simmel, G. 1902. *The sociology of Georg Simmel*, 1950 ed. Free Press, Glencoe, IL.
Simpson, W. 2001. The Quadratic Assignment Procedure (QAP). *North American Stata Users' Group Meetings*.
Singh, J., M.T. Hansen, J.M. Podolny. 2010. The World Is Not Small for Everyone: Inequity in Searching for Knowledge in Organizations. *Management Science* **56**(9) 1415-1438.
Soda, G., A. Zaheer. 2010. The Interplay of Formal and Informal Organizational Architecture: Implications for Performance. *Proceedings of the Academy of Management*.
Sorenson, O., T.E. Stuart. 2001. Syndication Networks and the Spatial Distribution of Venture Capital Investments. *American Journal of Sociology* **106**(6) 1546-1588.
South, S.J., C.M. Bonjean, W.T. Markham, J. Corder. 1982. Social Structure and Intergroup Interaction: Men and Women of the Federal Bureaucracy. *American Sociological Review* **47**(5) 587-599.
Srivastava, S.B., M.R. Banaji. 2011. Culture, Cognition, and Collaborative Networks in Organizations. *American Sociological Review* **76**(2) 207-233.
Stewart, J.Q. 1941. The "Gravitation," or Geographical Drawing Power, of a College. *Bulletin of the American Association of University Professors* **27**(1) 70-75.
Taylor, F.W. 1911. *Scientific Management*, 2nd Edition ed. Harper & Brothers, New York.
Thomas, D.A. 1990. The impact of race on managers' experiences of developmental relationships (mentoring and sponsorship): An intra-organizational study. *Journal of Organizational Behavior* **11**(6) 479-492.
Thompson, J.D. 1967. *Organizations in action; social science bases of administrative theory*, 2nd ed. Transaction Publishers, New Brunswick, NJ.
Tsai, W. 2001. Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. *Academy of Management Journal* **44**(5) 996-1004.
Turner, A.N., P.R. Lawrence. 1965. *Industrial jobs and the worker; an investigation of response to task attributes*. Harvard University, Division of Research, Graduate School of Business Administration, Boston,.
Tushman, M.L. 1977. Special Boundary Roles in the Innovation Process. *Administrative Science Quarterly* **22**(4) 587-605.
Tushman, M.L., D.A. Nadler. 1978. Information-Processing as an Integrating Concept in Organizational Design. *Academy of Management Review* **3**(3) 613-624.
Tzabbar, D., A. Vestal, R. Porter, S. McMahon. 2010. The Interdependencies of Formal and Informal Structure and the Exploration of New Technological Opportunities Among Geographically Dispersed Firms.
Williamson, O.E. 1975. *Markets and hierarchies, analysis and antitrust implications : a study in the economics of internal organization*. Free Press, New York.
Wooldridge, J.M. 1997. Quasi-Likelihood Methods for Count Data. *Handbook of Applied Econometrics* **2** 252-406.
Yakubovich, V., S. Shekshnia. 2009. The Formal Structure of Informal Networks: The Case of a Russian Corporate Bureaucracy.
Zipf, G.K. 1949. *Human behavior and the principle of least effort; an introduction to human ecology*. Addison-Wesley Press, Cambridge, MA.

Table 1
Descriptive Statistics and Correlations (N = 559,902 dyads)

		Mean	S.D.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
(1)	**Freq**	3.869	18.08	1																			
(2)	**SameOffice**	0.060	0.238	0.14	1																		
(3)	**SameBU**	0.319	0.466	0.19	0.18	1																	
(4)	**SameFunction**	0.419	0.493	0.13	0.14	0.40	1																
(5)	**SameSubfunction**	0.158	0.364	0.16	0.18	0.36	0.51	1															
(6)	**ln(Distance)**	6.140	1.970	-0.14	-0.79	-0.18	-0.13	-0.16	1														
(7)	**SameBand**	0.265	0.441	0.03	0.04	0.06	0.04	0.11	-0.04	1													
(8)	**SameGender**	0.581	0.493	0.01	0.02	0.01	0.04	0.03	-0.02	0.02	1												
(9)	**BothMale**	0.472	0.499	-0.01	0.02	-0.02	0.05	0.00	-0.02	0.02	0.80	1											
(10)	**BothFemale**	0.109	0.312	0.03	0.00	0.04	-0.01	0.04	-0.01	0.00	0.30	-0.33	1										
(11)	**log(InSample)**	6.466	0.829	0.14	0.04	0.14	-0.02	0.06	-0.07	0.03	-0.06	-0.13	0.11	1									
(12)	**log(OutOfSample)**	7.171	0.676	0.09	0.03	0.04	-0.08	-0.02	-0.05	0.03	-0.05	-0.11	0.09	0.60	1								
(13)	**log(AvgTenure)**	2.386	0.686	0.00	0.02	-0.05	-0.11	-0.08	-0.03	0.05	-0.01	-0.03	0.02	0.11	0.15	1							
(14)	**log(AvgBUSize)**	15.30	1.785	0.02	-0.05	0.29	0.10	0.03	0.01	-0.02	-0.03	-0.06	0.05	0.08	-0.07	-0.09	1						
(15)	**log(AvgFunctionSize)**	16.06	1.737	-0.04	-0.05	-0.02	0.36	0.00	0.08	-0.03	0.07	0.15	-0.14	-0.24	-0.26	-0.20	0.01	1					
(16)	**log(AvgSubfunctionSize)**	12.95	1.661	-0.01	-0.01	0.00	0.25	0.18	0.04	0.01	0.03	0.07	-0.07	-0.13	-0.25	-0.26	-0.05	0.62	1				
(17)	**log(AvgOfficeSize)**	10.28	1.613	0.01	0.14	0.01	-0.03	0.03	-0.18	0.00	-0.01	-0.02	0.02	0.08	0.06	0.09	-0.01	-0.13	-0.05	1			
(18)	**log(AvgBandSize)**	16.22	1.190	0.00	-0.01	0.00	-0.02	0.05	0.03	0.24	0.00	0.00	0.00	-0.01	-0.01	0.09	-0.03	-0.01	0.17	-0.05	1		
(19)	**OfficePctWomen**	36.41	7.016	0.01	-0.02	0.03	-0.03	0.00	-0.04	-0.01	-0.06	-0.11	0.08	0.14	0.10	0.02	0.12	-0.20	-0.16	0.26	-0.08	1	
(20)	**BUPctWomen**	36.85	8.933	-0.01	0.00	-0.01	-0.13	-0.03	-0.03	-0.01	-0.05	-0.11	0.09	0.08	0.10	0.12	0.38	-0.35	-0.28	0.12	-0.02	0.14	1
(21)	**SubfunctionPctWomen**	30.05	9.203	0.05	0.00	0.12	-0.03	0.10	-0.02	0.00	-0.10	-0.24	0.22	0.23	0.16	0.11	0.20	-0.47	-0.25	0.08	0.01	0.23	0.29

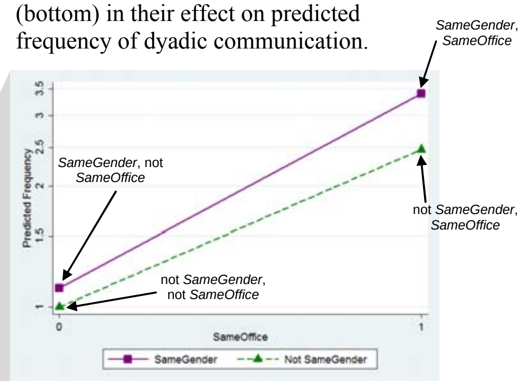
Table 2
Individual-level descriptive e-mail statistics, by gender. The final six columns indicate the gender distribution of the average man's and the average woman's contacts, both within and across business unit, office and functional boundaries.

Gender	Total Volume	Average Degree	Avg. Dyadic Frequency	Contacts' Gender Distribution	Gender Distribution of Contacts		Gender Distribution of Contacts		Gender Distribution of Contacts	
					Same BU	Across BUs	Same Office	Across Offices	Same Function	Across Functions
Male	2,564	38.2	12.4	M: 69.9% F: 30.1%	M: 70.6% F: 29.4%	M: 67.1% F: 32.9%	M: 72.0% F: 28.0%	M: 68.8% F: 31.2%	M: 71.7% F: 28.3%	M: 62.0% F: 38.0%
Female	2,850	46.4	13.3	M: 57.8% F: 42.2%	M: 57.3% F: 42.7%	M: 59.4% F: 40.6%	M: 58.8% F: 41.2%	M: 57.7% F: 42.3%	M: 57.1% F: 42.9%	M: 55.2% F: 44.8%

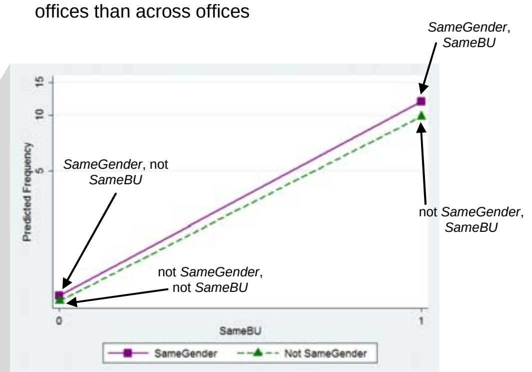
Table 3
Poisson quasi-maximum likelihood models of frequency of dyadic communication, including *SameGender* covariates and interactions.

	Baseline	Interactions with SameGender		
	(1)	(2)	(3)	(4)
SameOffice	1.032	0.904	1.033	1.032
	(0.077)**	(0.102)**	(0.077)**	(0.077)**
SameBU	2.352	2.351	2.285	2.352
	(0.040)**	(0.040)**	(0.047)**	(0.040)**
SameFunction	1.019	1.017	1.019	1.029
	(0.038)**	(0.038)**	(0.038)**	(0.046)**
SameSubfunction	1.198	1.203	1.199	1.198
	(0.037)**	(0.037)**	(0.037)**	(0.037)**
log(Distance)	-0.186	-0.186	-0.186	-0.186
	(0.010)**	(0.010)**	(0.010)**	(0.010)**
SameBand	0.297	0.293	0.296	0.298
	(0.029)**	(0.028)**	(0.029)**	(0.029)**
SameGender	0.155	0.111	0.071	0.167
	(0.027)**	(0.022)**	(0.027)**	(0.029)**
log(EmailVolume)	1.546	1.547	1.546	1.546
	(0.022)**	(0.022)**	(0.022)**	(0.022)**
log(prodBUSize)	-0.285	-0.285	-0.285	-0.285
	(0.010)**	(0.010)**	(0.010)**	(0.010)**
log(prodFunctionSize)	-0.016	-0.016	-0.016	-0.017
	(0.012)	(0.012)	(0.012)	(0.012)
log(prodSubfunctionSize)	-0.170	-0.171	-0.170	-0.170
	(0.010)**	(0.010)**	(0.010)**	(0.010)**
log(prodOfficeSize)	-0.097	-0.097	-0.097	-0.097
	(0.008)**	(0.008)**	(0.008)**	(0.008)**
OfficePctWomen	0.001	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)
BUPctWomen	-0.002	-0.003	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)
SubfunctionPctWomen	0.009	0.009	0.009	0.009
	(0.002)**	(0.002)**	(0.002)**	(0.002)**
SameOffice×SameGender		0.209		
		(0.099)*		
SameBU×SameGender			0.114	
			(0.044)**	
SameFunction×SameGender				-0.017
				(0.044)
Constant	-7.621	-7.604	-7.567	-7.627
	(0.271)**	(0.267)**	(0.270)**	(0.271)**
Observations	559,902	559,902	559,902	559,902

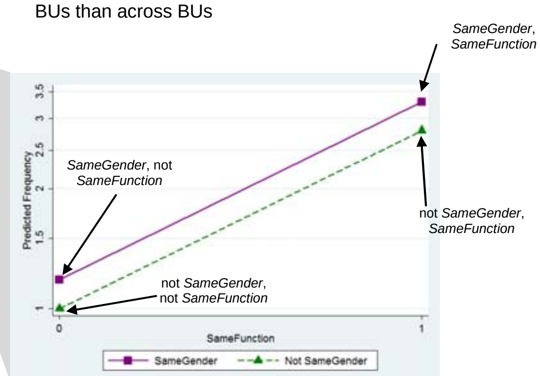
Table 3 Figures. Charts of the interaction of *SameGender* with *SameOffice* (top), *SameBU* (middle) and *SameFunction* (bottom) in their effect on predicted frequency of dyadic communication.



The slope of *SameOffice* is significantly larger in same gender dyads than in male-female dyads; the *SameGender* effect is larger within offices than across offices



The slope of *SameBU* is significantly larger in same gender dyads than in male-female dyads; the *SameGender* effect is larger within BUs than across BUs

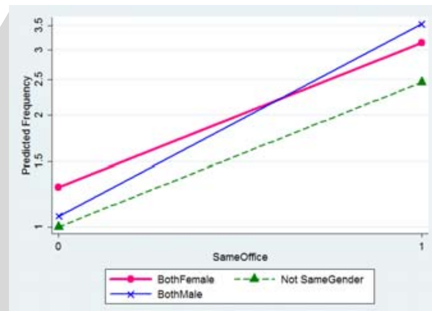


There is no significant difference in the slope of *SameFunction* in *SameGender* dyads compared to male-female dyads; the *SameGender* effect is the same within job functions as across job functions

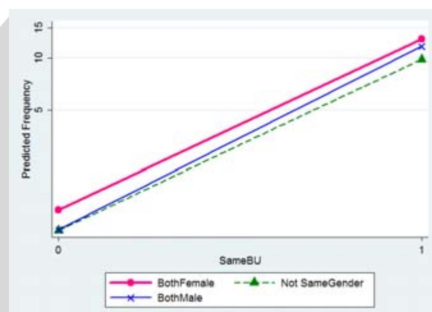
Table 4
Poisson quasi-maximum likelihood models of frequency of dyadic communication, including *BothMale* and *BothFemale* covariates and interactions.

	Baseline	BothMale/BothFemale Interactions		
	(5)	(6)	(7)	(8)
SameOffice	1.028	0.898	1.027	1.029
	(0.077)**	(0.102)**	(0.077)**	(0.077)**
SameBU	2.352	2.348	2.286	2.354
	(0.040)**	(0.040)**	(0.047)**	(0.040)**
SameFunction	1.019	1.018	1.021	1.024
	(0.038)**	(0.038)**	(0.038)**	(0.046)**
SameSubfunction	1.199	1.200	1.199	1.197
	(0.037)**	(0.037)**	(0.037)**	(0.037)**
log(Distance)	-0.186	-0.187	-0.186	-0.186
	(0.010)**	(0.010)**	(0.010)**	(0.010)**
SameBand	0.298	0.293	0.296	0.299
	(0.029)**	(0.028)**	(0.029)**	(0.029)**
log(EmailVolume)	1.543	1.540	1.543	1.542
	(0.022)**	(0.022)**	(0.022)**	(0.022)**
log(prodBUSize)	-0.285	-0.284	-0.285	-0.285
	(0.010)**	(0.010)**	(0.010)**	(0.010)**
log(prodFunctionSize)	-0.016	-0.014	-0.017	-0.016
	(0.012)	(0.011)	(0.012)	(0.012)
log(prodSubfunctionSize)	-0.170	-0.171	-0.170	-0.170
	(0.010)**	(0.010)**	(0.010)**	(0.010)**
log(prodOfficeSize)	-0.096	-0.096	-0.096	-0.096
	(0.008)**	(0.008)**	(0.008)**	(0.008)**
OfficePctWomen	0.001	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)
BUPctWomen	-0.002	-0.003	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)
SubfunctionPctWomen	0.008	0.009	0.009	0.008
	(0.002)**	(0.002)**	(0.002)**	(0.002)**
BothMale	0.129	0.065	0.011	0.157
	(0.028)**	(0.024)**	(0.031)	(0.033)**
BothFemale	0.230	0.247	0.272	0.184
	(0.042)**	(0.036)**	(0.042)**	(0.047)**
SameOffice×BothMale		0.300		
		(0.103)**		
SameOffice×BothFemale		-0.118		
		(0.151)		
SameBU×BothMale			0.162	
			(0.049)**	
SameBU×BothFemale			-0.056	
			(0.067)	
SameFunction×BothMale				-0.038
				(0.049)
SameFunction×BothFemale				0.067
				(0.072)
Constant	-7.568	-7.570	-7.506	-7.573
	(0.270)**	(0.265)**	(0.269)**	(0.270)**
Observations	559,902	559,902	559,902	559,902

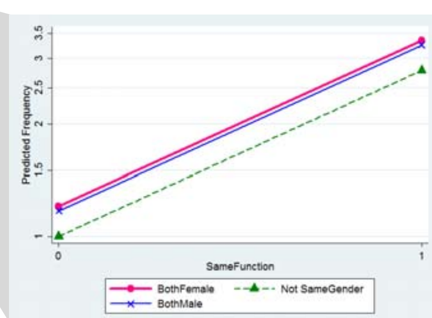
Table 4 Figures. Charts of the interaction of *BothMale* and *BothFemale* with *SameOffice* (top), *SameBU* (middle) and *SameFunction* (bottom) in their effect on frequency of dyadic communication.



The slope of *SameOffice* is significantly larger in *BothMale* dyads than in *BothFemale* or male-female dyads; the *BothMale* effect is larger within offices than across offices, but the *BothFemale* effect is not



The slope of *SameBU* is significantly larger in *BothMale* dyads than in *BothFemale* or male-female dyads; the *BothMale* effect is larger within BUs than across BUs, but the *BothFemale* effect is not



There is no significant difference in the slope of *SameFunction* in *BothMale* dyads, *BothFemale* or male-female dyads; both the *BothMale* and *BothFemale* effects are equal in magnitude within, compared to across, job functions